Supplementary Tables

Supplementary Table 1. Description of analytic cohorts

Cohort	Cohort description	All members genotyped	Subset with exome sequence for both parents and proband	Subset with exome sequence for both parents, proband and sibling
SSC	SSC quads	2,091	1,896 (90.7%)	1,729 (82.7%)
	European ancestry SSC quads	1,509	1,366 (90.5%)	1,221 (80.9%)
	SSC trios	493	449 (91.1%)	NA
	European ancestry SSC trios	342	311 (90.9%)	NA
PGC ASD	PGC ASD trios	3,870	NA	NA
	European ancestry PGC ASD trios	3,209	NA	NA

Ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***); trio families included both parents and the proband; quad families included both parents, the proband, and an unaffected sibling; count of families with all members genotyped refers to those remaining after imputation and data cleaning using the Ricopili pipeline[1].

Supplementary Table 2. Description of PGC ASD subcohorts

PGC ASD Cohort	Number of genotyped trios	Probands of European ancestry (%)	Reference for proband IQ measurement
Autism Center of Excellence, UCLA	215	67.4	Martin et al.[2]
Children's Hospital of Philadelphia	499	97.2	
Autism Genome Project (Group 1)	1,312	88.6	Anney et al.[3]
Autism Genome Project (Group 2)	942	77.3	
Johns Hopkins University	764	76.3	IQ measure not collected
Montreal/Boston Collection	138	76.1	
Total	3,870	82.9	

Ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***); see PGC Cross Disorders 2013 for more details about cohorts[4]; count of trios refers to those with all members remaining after imputation and data cleaning in the Ricopili pipeline[1].

Supplementary Table 3. Description of summary statistics from genome-wide association studies

Phenotype	Data source	Discovery sample size	P-value threshold for SNP inclusion in PRS	SNPs included in PGC ASD + SSC PRS	SNPs included in SSC-only PRS	Availability of summary statistics
Autism spectrum disorder (ASD)	iPSYCH-Broad Autism[5]	7,783 Cases, 11,359 Controls	0.1	25,634	28,552	Pre-publication consortium data; email mjdaly@atgu.mgh.harvard.edu or anders@biomed.au.dk for more information
Educational attainment (EA)	Discovery and replication meta-analysis, 23andMe excluded[6]	328,917 Individuals	1	155,210	262,851	Summary statistics publically available from Okbay et al.
Schizophrenia (SCZ)	PGC 2014[7]	36,989 Cases, 113,075 Controls	0.05	23,516	24,808	Summary statistics publically available from the Schizophrenia Working Group of the Psychiatric Genomics Consortium
Body mass index (BMI)	GIANT Consortium[8] (European meta-analysis)	322,154 Individuals	0.2	32,492	NA	Summary statistics publically available from the GIANT Consortium

SNPs excluded from PRS with SSC/PGC ASD imputation info score < 0.6; see **Online Methods:** ***Polygenic Risk Scoring*** for overview of summary statistics in the context of polygenic risk scoring.

Supplementary Table 4. Estimates of assortative mating from polygenic risk scores in SSC and PGC ASD

Cohort (European ancestry)	ASD PRS (r, p-value)	EA PRS (r, p-value)	SCZ PRS (r, p-value)
SSC	0.063 (7.13E-03)	-0.016 (0.49)	-0.0018 (0.94)
PGC ASD	0.0050 (0.78)	0.067 (1.36E-04)	0.037 (0.036)

R-values are Pearson correlation coefficients between maternal and paternal PRS within either SSC (n = 1,851 families) or PGC ASD (n = 3,209 families) for a given PRS with the first 10 principal components of ancestry regressed out (e.g., correlation between SSC mother ASD PRS and SSC father ASD PRS); p-values are the probability that there is no correlation between paternal and maternal PRS; analysis restricted to European ancestry families, with ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***).

Supplementary Table 5. Correlations between PRS in SSC and PGC ASD

		SSC		PGC ASD	
		ASD PRS	EA PRS	ASD PRS	EA PRS
Probands	EA PRS	r = 7.8E-03 P = 0.74	***n* = 1,852**	r = -0.010, P = 0.56	***n* = 3,209**
	SCZ PRS	r = 0.021, P = 0.36	r = 0.015, P = 0.53	r = 5.5E-03, P = 0.75	r = 0.036, P = 0.040
Mothers	EA PRS	r = 0.025, P = 0.28	***n* = 1,852**	r = -2.4E-03, P = 0.89	***n* = 3,209**
	SCZ PRS	r = 0.016, P = 0.50	r = 0.035, P = 0.13	r = 0.025, P = 0.15	r = 0.012, P = 0.48
Fathers	EA PRS	r = 0.050, P = 0.031	***n* = 1,851**	r = -0.014, P = 0.44	***n* = 3,209**
	SCZ PRS	r = 4.0E-04, P = 0.99	r = -1.5E-03, P = 0.52	r = -4.0E-03, P = 0.82	r = 0.015, P = 0.40

R-values are Pearson correlation coefficients; analysis performed in European ancestry cohorts of SSC and PGC ASD, with ancestry derived from analysis of genetic data (**Online Methods: *Sample Description***); first 10 principal components of ancestry regressed out of PRS (e.g., proband principal components of ancestry regressed out of proband PRS before analysis); p-values are the probability there is no correlation between the polygenic risk scores.

Supplementary Table 6. Parent PRS as a function of parent sex

			ASD PRS		EA PRS		SCZ PRS	
		n families (European ancestry)	Beta	p-value	Beta	p-value	Beta	p-value
Parent Sex (father = 1, mother = 0)	SSC	1,851	0.019	0.69	0.048	0.11	-0.076	0.021
	PGC ASD	3,209	-0.017	0.63	-0.0015	0.94	0.010	0.67

Effect size and p-values generated from three separate linear regressions predicting PRS from parent sex while controlling for the first 10 principal components of mother and father ancestry; p-value is the probability the means of the mothers' and fathers' PRS distributions are equal; ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***).

Supplementary Table 7. Mid-parent PRS as a function of proband sex

			Mid-parent ASD PRS		Mid-parent EA PRS		Mid-parent SCZ PRS	
		n families (European ancestry)	Beta	p-value	Beta	p-value	Beta	p-value
Proband Sex (male = 1, female = 0)	SSC	1,851	-0.019	0.79	0.057	0.17	-7.51E-03	0.87
	PGC ASD	3,209	0.034	0.49	8.52E-04	0.98	-0.012	0.72

Effect size and p-values generated from three separate linear regressions predicting mid-parent PRS from proband sex while controlling for 10 principal components of proband ancestry; p-value is the probability that mid-parent PRS does not differ by proband sex; ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***).

Supplementary Table 8. ASD Probands Over Inherit ASD Associated Polygenic Risk

Ancestry Restriction	Cohort	ASD PRS (pTDT deviation mean (SD), p-value)	EA PRS (pTDT deviation mean (SD), p-value)	SCZ PRS (pTDT deviation mean (SD), p-value)
None	SSC Probands (*n* = 2,584)	0.11 (0.96), 4.47E-09	0.074 (0.68), 1.32E-08	0.036 (0.37), 7.61E-07
	SSC Siblings (*n* = 2,091)	0.0086 (0.96), 0.68	-0.019 (0.63), 0.17	-0.011 (0.36), 0.16
	PGC ASD Probands (*n* = 3,870)	0.10 (1.02), 9.07E-10	0.076 (0.74), 1.86E-10	0.055 (0.52), 3.25E-11
	SSC+PGC ASD Probands (*n* = 6,454)	0.11 (0.99), 2.97E-17	0.075 (0.70), 1.41E-17	0.045 (0.44), 1.35E-16
European Ancestry	SSC Probands (*n* = 1,851)	0.11 (0.98), 2.27E-06	0.092 (0.86), 4.63E-06	0.081 (0.94), 1.94E-04
	SSC Siblings (*n* = 1,509)	-0.0074 (0.96), 0.76	-0.019 (0.81), 0.37	-0.037 (0.92), 0.12
	PGC ASD Probands (*n* = 3,209)	0.10 (1.03), 2.39E-08	0.088 (0.87), 1.11E-08	0.10 (0.92), 1.66E-10
	SSC+PGC ASD Probands (*n* = 5,060)	0.10 (1.01), 2.76E-13	0.089 (0.87), 2.40E-13	0.096 (0.93), 2.12E-13

P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test); ancestry derived from analysis of genetic data (**Online Methods: *Sample Description***).

Supplementary Table 9. BMI pTDT analysis

	BMI PRS (pTDT deviation mean (SD), p-value)
SSC Probands (*n* = 2,584)	8.2E-03 (0.38), 0.28
SSC Siblings (*n* = 2,091)	2.6E-03 (0.39), 0.76
PGC ASD Probands (*n* = 3,870)	0.013 (0.51), 0.12

P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test).

Supplementary Table 10. Comparison between pTDT deviation in discovery (SSC) and replication (PGC ASD) cohorts

	ASD PRS pTDT	EA PRS pTDT	SCZ PRS pTDT
No ancestry restriction	P = 0.67	P = 0.89	P = 0.08
European ancestry	P = 0.83	P = 0.89	P = 0.39

P-values were derived from two-sided, two-sample t-tests and reflect the probability that the means of the pTDT deviation distributions in **Supplementary Table 8** in SSC and PGC ASD are equal; number of subjects in each comparison available from **Supplementary Table 8**; ancestry derived from analysis of genetic data (**Online Methods:** ***Sample Description***).

Supplementary Table 11. pTDT analysis in probands with and without intellectual disability

Cohort	ID status	ASD PRS (pTDT deviation mean (SD), p-value)	EA PRS (pTDT deviation mean (SD), p-value)	SCZ PRS (pTDT deviation mean (SD), p-value)
SSC	With ID (*n* = 783)	0.13 (0.95), 9.15E-05	0.085 (0.62), 1.36E-04	0.037 (0.32), 1.20E-03
	Without ID (*n* = 1,795)	0.10 (0.97), 1.06E-05	0.068 (0.68), 2.24E-05	0.036 (0.40), 1.87E-04
PGC ASD	With ID (*n* = 558)	0.16 (1.08), 4.65E-04	0.13 (0.78), 9.46E-05	0.044 (0.53), 0.048
	Without ID (*n* = 948)	0.13 (0.98), 3.94E-05	0.12 (0.77), 1.71E-06	0.043 (0.64), 0.037
SSC + PGC ASD	With ID (*n* = 1,341)	0.14 (1.00), 1.50E-07	0.10 (0.67), 5.50E-08	0.038 (0.37), 1.69E-04
	Without ID (*n* = 2,743)	0.11 (0.97), 2.23E-09	0.084 (0.71), 5.02E-10	0.037 (0.45), 2.17E-05

P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test); ID (intellectual disability) = full-scale IQ < 70 (**Online Methods: *pTDT***).

Supplementary Table 12. Comparison between pTDT deviation in SSC + PGC ASD with and without ID

	ASD PRS pTDT	EA PRS pTDT	SCZ PRS pTDT
SSC + PGC ASD	P = 0.32	P = 0.47	P = 0.94

P-values were derived from two-sided, two sample t-tests and reflect the probability that the means of the pTDT deviation distributions in SSC + PGC ASD with (n = 1,341) and without ID (n = 2,743) are equal (**Supplementary Table 11)**.

Supplementary Table 13. pTDT in SSC probands with and without *de novo* mutations

	ASD PRS (pTDT deviation mean (SD), p-value)	EA PRS (pTDT deviation mean (SD), p-value)	SCZ PRS (pTDT deviation mean (SD), p-value)
Probands with CDNV (*n* = 221)	0.17 (1.04), 0.019	-0.012 (0.70), 0.80	0.054 (0.32), 0.013
Probands without CDNV (*n* = 2,124)	0.12 (0.95), 1.14E-08	0.087 (0.66), 2.02E-09	0.033 (0.37), 3.87E-05
ID probands without CNV or PTV (*n* = 533)	0.14 (0.93), 5.14E-04	0.10 (0.63), 1.64E-04	0.043 (0.30), 1.13E-03
Probands with constrained PTV and/or any CNV (*n* = 318)	0.18 (0.98), 1.66E-03	0.02 (0.69), 0.68	0.040 (0.36), 0.04
Probands without constrained PTV and/or any CNV (*n* = 2,028)	0.12 (0.96). 7.83E-08	0.087 (0.66), 4.35E-09	0.34 (0.36), 2.51E-05

See **Online Methods:** ***pTDT*** for CDNV definition; all variants are *de novo*; constrained PTV = protein truncating variant that was not observed in the publically available ExAC database and affected a gene with a high probability of being loss-of-function intolerant (pLI ≥ 0.9); P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test); ID (intellectual disability) = full-scale IQ < 70.

Supplementary Table 14. Prevalence of classes of *de novo* variation in SSC

Variant class	*n* cases with variant (% cases)	*n* controls with variant (% controls)	OR	p-value
Constrained deletions	57/2,587 (2.2%)	10/2,100 (0.5%)	4.71	3.17E-07
Large unconstrained deletions	8/2,587 (0.3%)	1/2,100 (0.05%)	6.51	0.048
Contributing CNV deletions	65/2,587 (2.5%)	11/2,100 (0.5%)	4.89	1.80E-08
Non-contributing deletions	43/2,587 (1.7%)	29/2,100 (1.4%)	1.21	0.48
Contributing PTVs	167/2,346 (7.1%)	37/1,736 (2.1%)	3.52	4.12E-14
Non-contributing PTVs	182/2,346 (7.8%)	120/1,736 (6.9%)	1.09	0.50
CDNVs	221/2,346 (9.4%)	45/1,736 (2.6%)	3.91	6.56E-20

CNV = copy number variant; PTV = protein truncating variant (frameshift, splice acceptor, splice donor, nonsense); constrained deletions = deletions containing a gene that was predicted to be intolerant of heterozygous loss of function variation (probability of being loss-of-function intolerant (pLI) ≥ 0.9); large unconstrained deletions = deletions ≥ 500 kb that do not contain a gene predicted to be intolerant of heterozygous loss of function variation; contributing CNV deletions = either 1) constrained deletion or 2) large unconstrained deletion; non-contributing deletions = *de novo* deletions that were neither constrained nor large unconstrained; contributing PTV = variant that was not observed in the publically available ExAC database and affected a gene with a high probability of being loss-of-function intolerant (pLI ≥ 0.9); non-contributing PTV = *de novo* PTV that is not contributing; CDNV = either contributing CNV deletion or contributing PTV; deletions roster was genotyped SSC probands and siblings; PTV and CDNV roster was genotyped and sequenced SSC probands and siblings; OR = odds ratio from case-control Fisher's Exact test; p-values generated from Fisher's Exact test indicate probability that the variant class is equally likely to be seen in cases and controls.

Supplementary Table 15. Relationship between rate of CDNVs and adverse co-occurring neurodevelopmental outcomes

	Motor delay	Seizures	ID	No co-occurring outcomes
Number of probands in category (SSC prevalence of phenotype)	147 (6.3%)	183 (7.8%)	705 (30.2%)	1,476
Count of CDNVs	38	33	91	111
CDNV rate (p-value controlling for proband sex)	0.26 (2.57E-08)	0.18 (3.8E-04)	0.13 (6.4E-03)	0.075

CDNV rate was calculated by dividing the count of CDNVs in a category by the number of probands in the category; p-values were from Poisson regression predicting CDNV count from present/absence of each co-occurring neurodevelopmental outcome and proband sex, and estimate the probability that the rate of CDNVs in a co-occurring outcome category was equal to the rate in probands with no co-occurring outcomes; motor delay was walking unaided at or after 19 months after birth; ID = intellectual disability = full-scale IQ < 70; CDNVs = contributing *de novo* variants (**Online Methods: De novo *variant analyses***); analytic cohort included SSC probands who were both genotyped and sequenced.

Supplementary Table 16. CDNV male:female ratio grouped by adverse co-occurring neurodevelopmental outcomes

Count of co-occurring neurodevelopmental outcomes in probands	Number of probands in co-occurring neurodevelopmental category (Number of probands in category with CDNV)	Proband CDNV rate (OR, p-value[a])	Observed male:female proband CDNV carrier ratio	(Expected male:female proband ratio) / (observed male:female proband CDNV carrier ratio) (p-value[b])
0	1,476 (105)	0.075 (3.15, 3.88E-10)	4.53	1.42 (0.095)
1	719 (77)	0.11 (4.53 (5.53E-15)	2.67	2.41 (6.81E-04)
2	134 (33)	0.25 (10.18, 6.94E-23)	2.00	3.21 (1.57E-03)
3	16 (6)	0.38 (15.05, 9.08E-10)	0.50	12.84 (2.87E-03)

Co-occurring neurodevelopmental outcomes (delayed walking, intellectual disability, seizures) are described in **Online Methods: De novo *variant analyses***, as are CDNVs; observed male:female ratio is the ratio of male:female CDNV carriers within a given outcome category; proband CDNV rate is the count of CDNVs in probands in each outcome category divided by count of probands in the outcome category; odds ratio (OR) was calculated from Poisson regression predicting CDNV count from case/control status for all controls (n = 1,736) and cases in the outcome category, controlling for maternal and paternal age at birth of the child; [a]p-value is derived from the Poisson regression and estimates the probability that the proband CDNV rate is equal to the control CDNV rate (CDNV rate in 1,736 SSC controls is 0.024 variants/exome); [b]p-value was calculated using a Fisher's exact test and estimates probability that the expected male:female proband ratio in SSC (overall SSC ratio of probands who were both genotyped and sequenced, (6.42)) was equal to the observed male:female proband ratio of CDNV carriers in the co-occurring neurodevelopmental outcome category.

Supplementary Table 17. pTDT as a function of adverse co-occurring neurodevelopmental outcomes

Adverse co-occurring neurodevelopmental outcome category	Probands in category	ASD PRS (pTDT deviation mean (SD), p-value)	EA PRS (pTDT deviation mean (SD), p-value)	SCZ PRS (pTDT deviation mean (SD), p-value)
0	1,475	0.12 (0.96), 3.00E-06	0.076 (0.68), 2.02E-05	0.029 (0.41), 5.75E-03
≥ 1	869	0.13 (0.97), 4.75E-05	0.081 (0.65), 2.59E-04	0.044 (0.32), 4.58E-05

Co-occurring neurodevelopmental outcomes (delayed walking, intellectual disability, seizures) are described in **Online Methods: De novo *variant analyses***; analytic cohort is genotyped and sequenced SSC probands with pTDT available; P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test).

Supplementary Table 18. Distinct polygenic risk factors are independently over transmitted to ASD probands

Cohort	ASD PRS (beta, p-value)	EA PRS (beta, p-value)	SCZ PRS (beta, p-value)
SSC Probands (*n* = 2,584)	0.072, 1.28E-03	0.113, 6.57E-04	0.040, 1.89E-02
PGC ASD Probands (*n* = 3,870)	0.065, 4.08E-04	0.109, 1.00E-04	0.052, 2.44E-03
SSC+PGC ASD Probands (*n* = 6,454)	0.068, 1.79E-06	0.112, 1.76E-07	0.046, 1.38E-04

For each of the three cohorts, we performed a single logistic regression predicting proband (1) or mid-parent (0) status from each of the three PRS; p-values estimate the probability with which the mid-parent and proband means are equal, controlling for the other two PRS.

Supplementary Table 19. IQ effect of ASD-associated genetic risk factors

	ASD PRS	EA PRS	SCZ PRS	CDNV
Genetic correlation with IQ in the general population (r, p-value), from Hagenaars et al. [9]	0.187, 0.005	0.720, 2.0E-38	-0.295, 3.5E-11	NA
ASD IQ associations (beta, p-value)	0.45, 0.31	1.39, 0.03	-1.84, 0.006	-10.26, 1.45E-05

The three ASD IQ - PRS associations are from three linear regressions predicting full-scale proband IQ from each PRS, controlling for proband sex and the first 10 principal components of proband ancestry; the ASD IQ – CDNV association is from a linear regression predicting full-scale proband IQ from CDNV presence/absence, controlling for proband sex; all four associations were performed in genotyped and sequenced European ancestry Simons Simplex Collection probands (n = 1,674); p-values estimate probability of no association between genetic factor (polygenic risk or genetic correlation) and IQ; genetic correlation results from Hagenaars et al.[9]

Supplementary Table 20. Sibling-based pTDT is less statistically powered than parent-based pTDT

	ASD PRS (pTDT deviation mean (SD), p-value)	EA PRS (pTDT deviation mean (SD), p-value)	SCZ PRS (pTDT deviation mean (SD), p-value)
SSC Sibling pTDT	0.084 (1.022), 1.81E-04	0.084 (0.769), 6.06E-07	0.047 (0.470), 5.54E-06
SSC Parent pTDT	0.135 (0.978), 3.83E-10	0.073 (0.662), 5.01E-07	0.040 (0.361), 3.54E-07

Sibling and parent comparisons performed with same cohort (n = 2,091 quads) to facilitate comparison; P-values denote the probability that the mean of the pTDT deviation distribution is 0 (two-sided, one-sample t-test).

Supplementary Table 21. CNV analysis integrating parental age

y	x1	No parental age controls (n = 2,587 cases) (n = 2,100 controls)		Controlling for paternal and maternal age at birth of child (n = 2,346 cases) (n = 1,761 controls)	
		OR	**p-value**	**OR**	**p-value**
Case status (proband = 1, 0 = unaffected sibling)	Constrained and/or ≥ 500 kb CNV deletion	4.89	1.22E-06	5.35	1.01E-05
	Constrained CNV deletion	4.71	6.69E-06	5.08	2.07E-05
	Unconstrained ≥ 500 kb CNV deletion	6.51	0.077	NA	NA
	Unconstrained and < 500 kb CNV deletion	1.21	0.44	1.23	0.46
	CNV duplication with gene	3.11	3.60E-05	2.99	1.70E-04
	CNV duplication without a gene	2.85	0.19	2.63	0.23

All variants are *de novo*; constrained refers to CNVs containing genes that are intolerant of heterozygous loss of function variation (probability of being loss-of-function intolerant (pLI) ≥ 0.9); NA denotes an analysis where all control carriers were missing parental age data; analytic model was logistic regression predicting proband/control status from count of CNVs; ORs are interpreted as the increased likelihood of proband status given presence of a variant; p-values test the null hypothesis that the OR is equal to 1 (no association between variant and case status).

Supplementary Table 22. Relationship between proband IQ, CDNV status and mid-parent PRS

	Mid-parent ASD PRS (beta, p-value)	Mid-parent EA PRS (beta, p-value)	Mid-parent SCZ PRS (beta, p-value)
Proband CDNV status (*n* = 1,677)	0.081, 0.38	0.026, 0.68	-6.99E-04, 0.99
Proband FSIQ (controlling CDNV status) (*n* = 1,674)	0.0022, 0.025	0.0017, 0.014	-0.0016, 0.0096

For each polygenic risk category, we performed a linear regression predicting mid-parent PRS from proband CDNV status (presence = 1, absence = 0), and next a linear regression predicting mid-parent PRS from proband full-scale IQ and CDNV status; in the first row, p-values indicate the probability of no association between CDNV status and mid-parent PRS; in the second row, p-values indicate the probability of no association between proband FSIQ and mid-parent PRS, controlling for proband CDNV status; analytic cohort is Simons Simplex Collection European ancestry families with probands who were both genotyped and sequenced and genotyped parents; FSIQ = full-scale IQ.

Supplementary Notes

Glossary

ASD Proband An individual diagnosed with autism spectrum disorder
ASD PRS Polygenic Risk Score for Autism Spectrum Disorder
BMI PRS Polygenic Risk Score for Body Mass Index
CDNV Contributing *de novo* variant
CNV Copy number variant (deletion or duplication)
EA PRS Polygenic Risk Score for Educational Attainment
iPSYCH-Broad Autism Group The Lundbeck Foundation Initiative for Integrative Psychiatric Research-Broad Institute Autism Group
Mid-Parent PRS Average polygenic risk score of mother and father in a given family
PGC ASD Psychiatric Genomics Consortium Autism Group
pLI Probability of being Loss-of-Function Intolerant (probability that gene is intolerant of heterozygous loss-of-function variation)[10]
PRS Polygenic Risk Score
pTDT Polygenic transmission disequilibrium test
PTV Protein-truncating variant (frameshift, splice acceptor, splice donor, nonsense)
SCZ PRS Polygenic Risk Score for Schizophrenia
SSC Simons Simplex Collection

Supplementary Note 1: Estimates of assortative mating on polygenic risk. We examined evidence for assortative mating by correlating maternal and paternal polygenic risk for each trait in both SSC and PGC ASD. We restricted our analysis to families of European ancestry (**Online Methods:** ***Sample Description***) to avoid ancestral confounding. For each of ASD, EA and SCZ PRS, we regressed out the first 10 principal components of ancestry of each parent and correlated the residuals between mothers and fathers (**Supplementary Table 4**).

Supplementary Note 2: Parent PRS as a function of parent sex. We analyzed whether mothers and fathers differed with regard to average ASD, SCZ, or EA PRS. In each of European ancestry SSC and PGC ASD cohorts (**Online Methods:** ***Sample Description***), we performed three linear regressions, each predicting polygenic risk for ASD, EA and SCZ from parent sex while controlling for the first 10 principal components of each parent's ancestry (**Supplementary Table 6**).

Supplementary Note 3: Mid-Parent PRS as a function of proband sex. We analyzed whether mid-parent PRS differed as a function of proband sex (European ancestry only). In each of SSC and PGC ASD, we performed three linear regressions, predicting mid-parent polygenic risk for ASD, EA and SCZ from proband sex, controlling for the first 10 principal components of proband ancestry (**Supplementary Table 7**).

Supplementary Note 4: European ancestry pTDT. We repeated pTDT for ASD, EA and SCZ PRS in the four European ancestry cohorts defined in the Online Methods: European SSC probands (n = 1,851), European SSC unaffected siblings (n = 1,509), European PGC probands (n = 3,209), and the combination of European SSC and European PGC probands (n = 5,060). Polygenic transmission before and after ancestry restriction was largely consistent (**Supplementary Table 8**). In the European ancestry pTDT, there were no significant differences in pTDT deviation across the polygenic risk categories ($P > 0.05$ for all comparisons) (**Supplementary Figure 2**). Without ancestry restriction (**Figure 1a**), the mean ASD and EA pTDT deviation values were significantly greater than the mean SCZ pTDT deviation value ($P < 0.005$ for SSC+PGC ASD proband cohort). The SCZ PRS is most stratified by ancestry (**Supplementary Figure 3**), and limiting the pTDT analysis to families of European ancestry reduced the variance of the mid-parent SCZ PRS distribution. As the pTDT deviation value was normalized by the standard deviation of the mid-parent PRS, the mean proband deviation from the mid-parent value appeared larger in the European ancestry SCZ analysis than without an ancestry filter ($P = 3.10E-04$); if the deviations were not standardized by mid-parent PRS, this difference was eliminated ($P = 0.89$). This suggested no difference in pTDT deviation effect sizes for ASD, EA and SCZ PRS after controlling for ancestral effects.

Supplementary Note 5: Body Mass Index Polygenic Risk. To identify the optimal p-value threshold for polygenic risk scoring, we first calculated polygenic risk for all of the SSC at the ten standard p-values thresholds (**Online Methods:** ***Polygenic Risk Scoring***). We then associated the resulting polygenic risk scores with reported body mass index (BMI) in the SSC ($BMI = Weight/Height^2$). To do so, we identified the European ancestry SSC subcohort (**Online Methods:** ***Sample Description***) and regressed out the first 10 principal components of ancestry from each individual's polygenic risk scores. We also removed SSC individuals with BMI at least 3 standard deviations from the cohort mean.. We then correlated the BMI of the

remaining individuals with their polygenic risk scores, and identified P = 0.2 as the threshold that resulted in the strongest association (r = 0.14, P = 8.68E-09).

Supplementary Note 6: Sibling pTDT. We performed a sibling-based pTDT analysis to compare its statistical power with that of the parent-based pTDT. Using genotyped SSC quads (*n* = 2,091), we calculated sibling pTDT deviation as follows:

$$pTDT\ deviation_{sibling} = \frac{PRS_{Proband} - PRS_{Sibling}}{SD(PRS_{Siblings})}$$

As this analysis was specific to SSC, we generated polygenic risk scores using info thresholds from SSC imputation. For comparison of statistical power, we performed parent pTDT in the same cohort of quads (**Supplementary Table 20**). The loss of statistical power in the sibling-based pTDT was due to the increased variance of the distribution of sibling PRS relative to mid-parent PRS. The mid-parent distribution has reduced variance due to averaging of parent values.

Supplementary Note 7: *De novo* duplications. Duplications of constrained genes (pLI ≥ 0.9) were not associated with ASD risk after controlling for duplication size and maternal age at birth of child (P = 0.92, logistic regression). However, pLI may not be a good indicator of genes that are sensitive to duplication. Duplications may result in gain of function, or in other changes that do not result in loss of function, but increase risk for ASDs.

Supplementary Note 8: Association between co-occurring neurodevelopmental outcomes and proband sex ratio. We first calculated an expectation for the male:female ratio (6.42) as the total count of male SSC probands (*n* = 2,029) over the total count of female SSC probands (*n* = 316) (cohort: SSC probands both sequenced and genotyped). Next, we calculated the observed male:female ratio of CDNV carriers within each of the four categories of co-occurring neurodevelopmental outcomes (**Supplementary Table 16**). We used Fisher's exact tests to determine the significance of the difference between expected and observed male:female ratio in each category. The observed male:female ratio was significantly lower than expected for probands with at least one co-occurring neurodevelopmental phenotype (P < 5.00E-03).

Supplementary Note 9: pTDT in expanded set of CDNVs. We expanded our pTDT analysis by examining whether ASD associated risk was over inherited by carriers of a broader set of *de novo* mutations (cohort: genotyped and sequenced SSC, *n* = 2,346). Our expanded *de novo* mutation set included constrained PTVs (not observed in ExAC, pLI ≥ 0.9) and all *de novo* copy number variants (deletions and duplications) (*n* = 318 probands with ≥ 1 variant) (**Supplementary Table 13**). We also conducted pTDT in SSC genotyped and sequenced probands with full-scale IQ < 70 who lacked any *de novo* PTV or CNV (**Supplementary Table 13**).

Supplementary Note 10: Relationship between mid-parent PRS, proband IQ, and CDNV status (Supplementary Table 22). Next, we analyzed whether mid-parent PRS varied as a function of whether the proband carried a contributing *de novo* variant (CDNV, **Online**

Methods: De novo *variant analyses*). Our analytic cohort was genotyped and sequenced SSC families with a European ancestry-defined proband (n = 1,677). A subset of these families had a proband with at least one CDNV (n = 161 probands). We performed three linear regressions with CDNV status and first 10 proband principal components of ancestry as the independent variables, and mid-parent PRS for ASD, EA and SCZ as the dependent variables.

Finally, we analyzed whether mid-parent PRS varied as a function of proband IQ, controlling for whether the proband carried a CDNV. Our analytic cohort was genotyped and sequenced SSC families with a European ancestry-defined proband and IQ assessment available (full scale IQ) (n = 1,674). A subset of these families had a proband with at least one contributing *de novo* event (n = 161). We performed three linear regressions with proband IQ, CDNV status and first 10 proband principal components of ancestry as independent variables, and mid-parent PRS as the dependent variable.

References

1. Ricopili. Rapid Imputation Consortium Pipeline.
2. Martin, L.A. & Horriat, N.L. The effects of birth order and birth interval on the phenotypic expression of autism spectrum disorder. *PLoS One* **7**, e51049 (2012).
3. Anney, R. *et al.* Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* **21**, 4781-92 (2012).
4. Cross-Disorder Group of the Psychiatric Genomics, C. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984-94 (2013).
5. Robinson, E.B. *et al.* Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet* **48**, 552-5 (2016).
6. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-42 (2016).
7. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).
8. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).
9. Hagenaars, S.P. *et al.* Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Mol Psychiatry* (2016).
10. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-91 (2016).